

CS 4495 Computer Vision

Introduction

Aaron Bobick

School of Interactive Computing

Who are we?



Professor:

Aaron Bobick

afb@cc.gatech.edu

Office: CCB 316

Office hours:

Tues 1-2pm (email is **much** better)

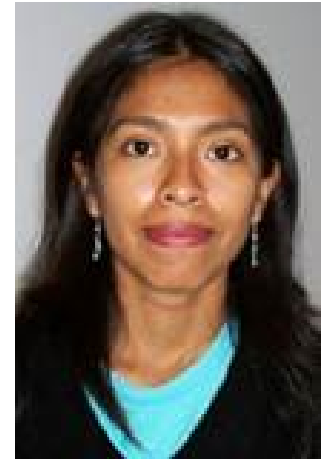
TA:

Ana Huaman

ahuaman3@gatech.edu

Office: CCB 2xx

Office hours: TBD



Yin Li

happyharrycn@gmail.com

Office: CCB 308

Office hours: TBD



Where are you?

- **CS 4495 Computer Vision**

- **Web site:** <http://www.cc.gatech.edu/~afb/classes/CS4495-Fall2014/>
 - Will have posted calendar/syllabus with posted slides, problem sets with data, other administrative stuff.
- **T-square:** The usual stuff. There is/will be a web page under resources that points to the class web page listed above.
- **Slides:** PDFs will be posted by linking to the calendar. Hopefully draft by class time.
- **Piazza:** You will all receive an invitation to Piazza for CS4495. If not, send us email. This was invaluable for discussions for problem sets.
 - **Announcements** will likely be done through both T-Square (email) and Piazza
- **Matlab access:** if you don't know how to get Matlab access, first ask a friend. Then come see TAs or me. And ask about Python/OpenCV

Where are you?

- **CS 7495 Computer Vision**
- **All of the above plus...**
- **Web site:**
<http://www.cc.gatech.edu/~afb/classes/CS7495-Fall2014/>
 - It is not there yet!

Who are you? - CS4495

- From the web site:

This is a senior level undergraduate course for those interested in computer vision. The GR section is open to graduate students who are MSCS students or who need a solid undergraduate grounding before /while taking the graduate version (CS7495).

- What do you know?

- Data structures – you'll be writing code that builds representations of images, features, and geometric constructions
- Matlab (or Python with OpenCV or if really desperate C++/OpenCV)
- More math than most CS courses: Linear algebra, Vector Calculus, Linear algebra, Probability, Linear algebra
- No CV assumed but maybe you've hacked an image?

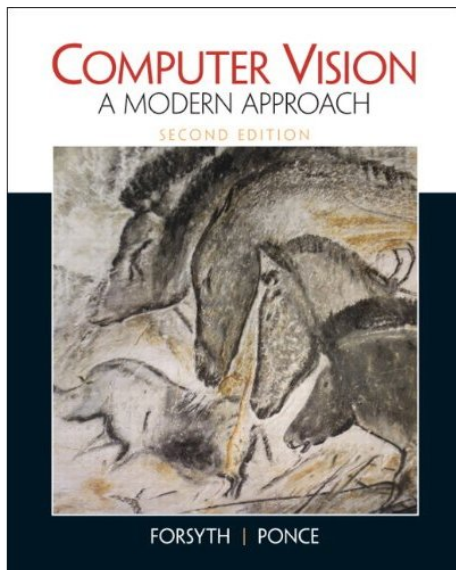
Who are you? - CS7495

- From the (soon to be) web site:
The graduate class in computer vision that assumes a general background at the undergraduate level.
- What do you know?
 - Everything before plus basic computer vision

What will you read?

Forsythe and Ponce:

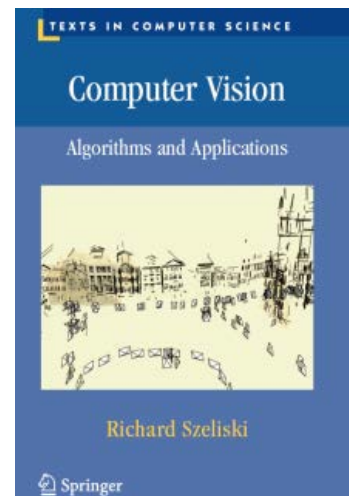
*Computer Vision: A
Modern Approach (2nd
Ed)*



(no pdf)

• Rick Szeliski's
book:

*Computer Vision:
Algorithms and
Applications*



(pdf)

What you will do? – CS4495

- The grade is mostly problem set based (85%). Currently 8 tentatively scheduled:
 - 0 Due Thurs Aug 29th midnight. *Almost already posted* This is just to make sure you can read, numerically manipulate, and display images. ***This one is no work. Others are serious work.***
 1. Edges, lines, Hough transform
 2. Stereo Matching
 3. Calibration – where is the camera given the points.
 4. Relative motion – SIFT features to find where the cameras are
 5. Motion flow – or maybe Pyramid match kernel????
 6. Tracking
 7. Motion History Images (action recognition)
- And there will be an exam – 15%

What you will do? – CS7495

- There will be a paper assigned per week of more current or advanced material.
- There will be an optional one hour meeting each week to discuss the paper (or projects).
 - There will be a simple summary question for each paper that you'll hand in. Easy to answer if you attend the session.
- The grade is mostly problem on three projects.
- See previous year's web site:
- <http://www.cc.gatech.edu/~dellaert/13F-7495/Syllabus.html>

When will you do them? – CS4495

- Except for PS0 0, they will all be due on Sunday at 11:59pm.
- Late submissions will only be accepted at full credit with prior approval. Otherwise 50% (yes half) reduction.
- TA/Prof ***not*** obligated to get back to you about permissionthe weekend it is due! At your own risk.
- I will be very adamant about this. Not fair to others or the TAs.

When will you do them? – CS7495

- Summary questions will be weekly.
- There will be 6 deadlines for the projects – each has a proposal and a final writeup.

With whom will you do it? – CS4495

- Honesty/Integrity policy (from web site):
Problem sets are to be done individually but you may collaborate at the “white board level” helping each other with algorithms and general computation, BUT YOUR CODE MUST BE YOUR OWN.
- Do not hand in other people’s code unless you (1) say you are, and (2) you want no credit for that section. We will be explicit about what previous or provided code you can use.
- Grad students (CS4495GR) will be graded as a separate pool so as to not penalize undergraduates. (Some day it will be CS6495.) Otherwise, deliverables are identical.

How will you do it? – CS4495

- We will mostly support Matlab/Octave
 - In class, most of the time we will use Matlab to handle the image processing (though we'll learn some of that) and then your data structures for the vision.
- If you want be even more “job ready”
 - OpenCV is a Python (and C/C++) package for computer vision (Linux or Windows)
 - Can be painful to install/use because OpenCV breaks with every version.
 - Grads: Other graduate students in RIM (Robotics and Intelligent Machines) can help.
 - TAs will not be obligated to help with OpenCV “issues” but we will make sure problem sets are Python-doable.
 - One prior student” “Don’t let them use OpenCV. Too hard to do the extra stuff like graphs etc. “

With whom will you do it? – CS7495

- Summaries on your own (duh).
- Each project must be done by 2 or 3 people. No more, no less. If you can not find a partner we'll help you pair up.

Any questions so far...

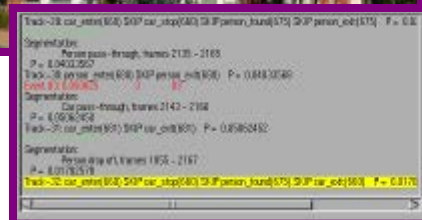
Who am I...

- Background: degrees in Math, CS but PhD in CogSci/AI; interest in high level perception and cognition
- Faculty at the MIT Media Lab for a while – developed lots of work in machine understanding of action from video
- Came here 15 years ago, and collected hats:
 - Was GVU director
 - Founding chair of the School of Interactive Computing
 - But the most fun is being professor in Computational Perception and Robotics!

Research I used to do...

Action recognition from video

- Lots of domains/levels of complexity:
 - Body motions
 - Gesture recognition,
 - Football plays
 - Aware Home
 - Surveillance

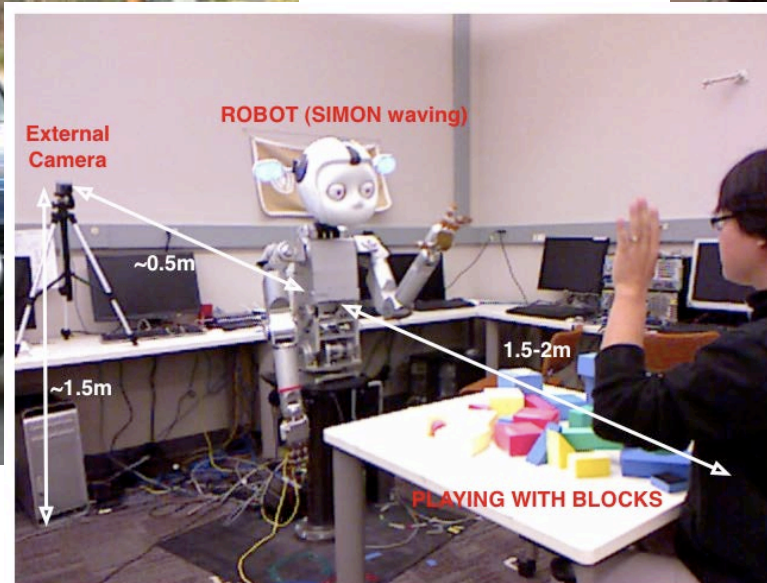


What I am (or will be) doing now...

Robots that “see”... not (just) a question of geometry
but *understanding*.



Outdoor



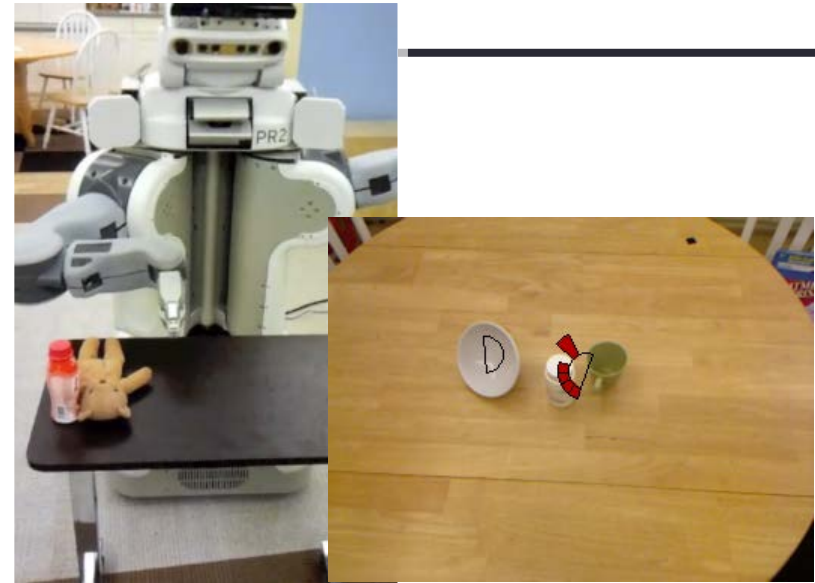
Human Robot Interaction



Indoor

Two specific projects:

NEW
(Vision for) Human-robot
collaboration in manufacturing
(with BMW!)



Affordance-based
perception: Robot learning
the “affordances” of objects
and how to use in planning
and acting.

Why study Computer Vision?

- Images (and movies) have become ubiquitous in both production and consumption.
- Therefore application to manipulate images (movies) are becoming core.
- As are systems that *extract information* from imagery
 - Surveillance
 - Building 3D representations
 - Motion capture assisted
- But most of all...

It is a really deep and cool set of problems!

What is computer vision?



Terminator 2

Every picture tells a story



- Goal of computer vision is to write computer programs that can interpret images

Can computers match (or beat) human vision?



- Yes and no (but mostly no!)
 - humans are much better at “hard” things
 - computers can be better at “easy” things

Human perception has its shortcomings...

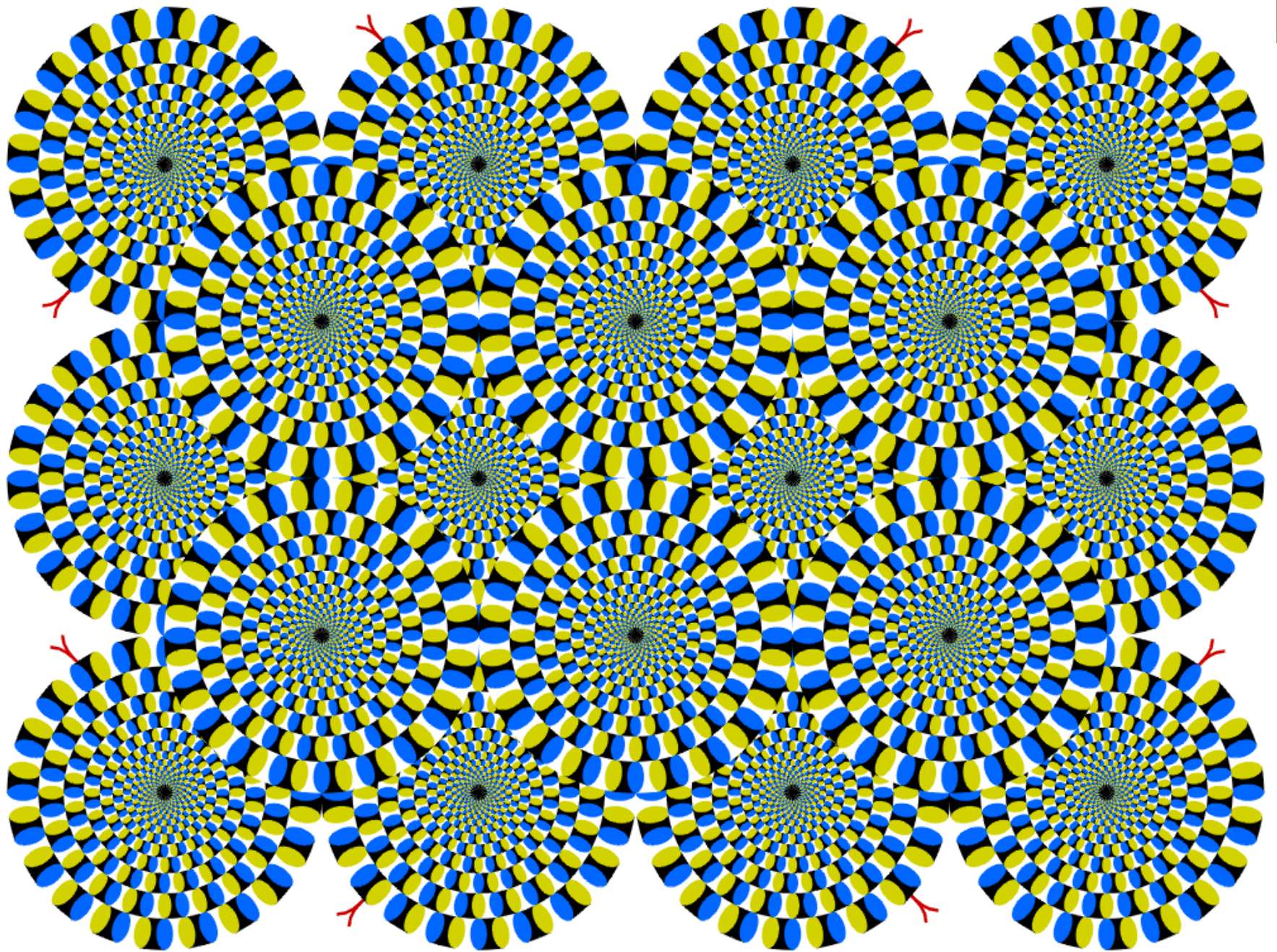


[Sinha and Poggio, *Nature*, 1996](#)

Human perception has its shortcomings...



[Sinha and Poggio, *Nature*, 1996](#)



Current state of the art

- The next slides show some examples of what current vision systems can do

Earth viewers (3D modeling)

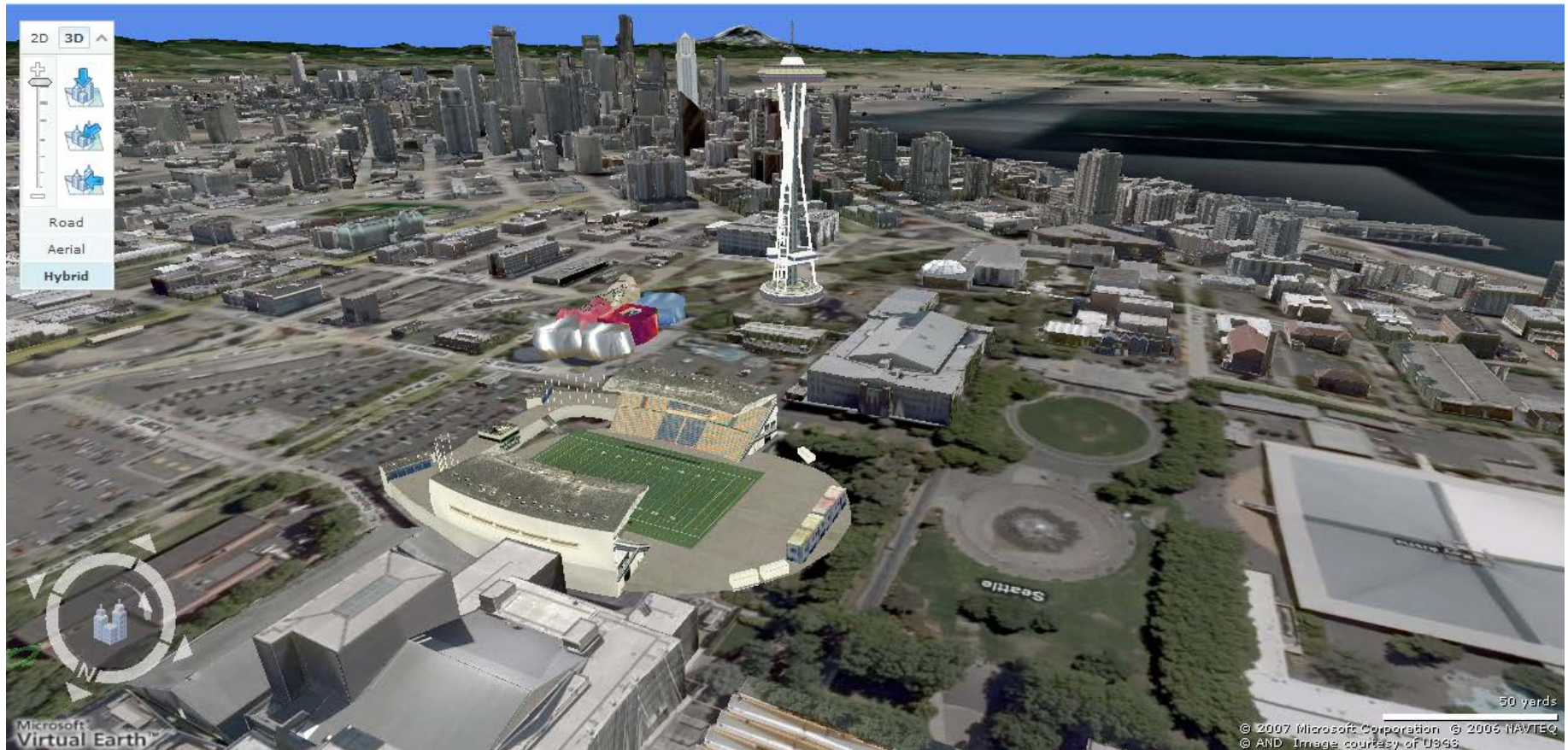


Image from Microsoft's [Virtual Earth](#)
(see also: [Google Earth](#))



- Home
- Try it
- What is Photosynth?
- Collections
- Team blog
- Videos
- System requirements
- About us
- FAQ



The **Photosynth Technology Preview** is a taste of the newest - and, we hope, most exciting - way to **view photos** on a computer. Our software takes a large collection of photos of a place or an object, analyzes them for similarities, and then displays the photos in a reconstructed **three-dimensional space**, showing you how each one relates to the next.

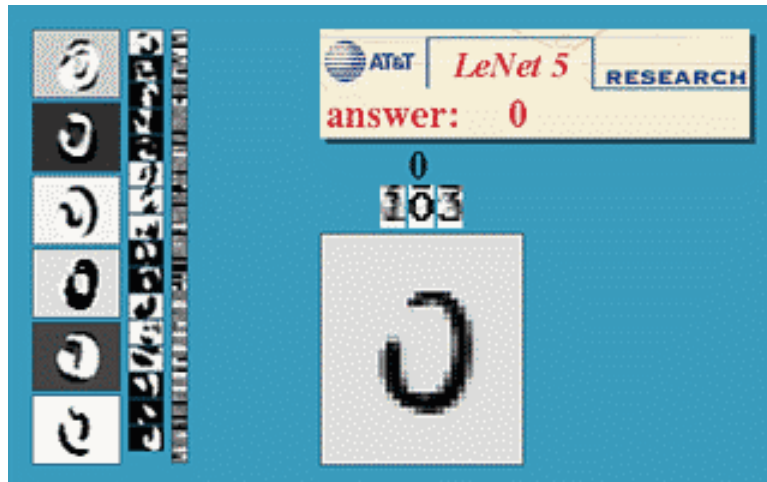
<http://labs.live.com/photosynth/>

Based on [Photo Tourism technology](#) developed here in UW.
by Noah Snavely, Steve Seitz, and Rick Szeliski

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs

<http://www.research.att.com/~yann/>



License plate readers

http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Face detection



- Many new digital cameras now detect faces
 - Canon, Sony, Fuji, ...

Smile detection?

The Smile Shutter flow

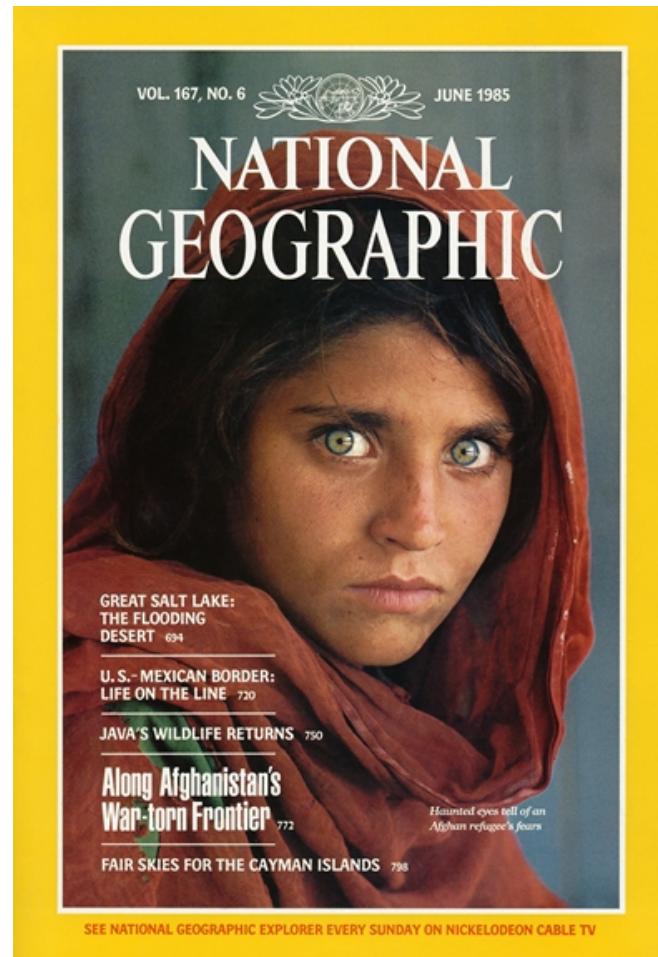
Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



[Sony Cyber-shot® T70 Digital Still Camera](#)

Steve Seitz

Face recognition



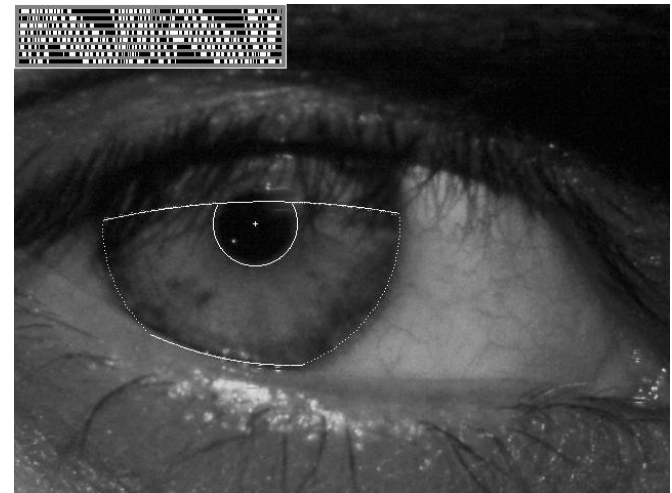
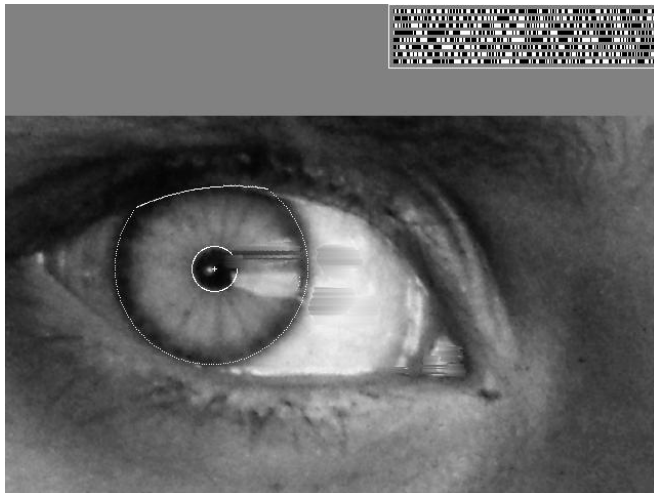
Who is she?

Steve Seitz

Vision-based biometrics



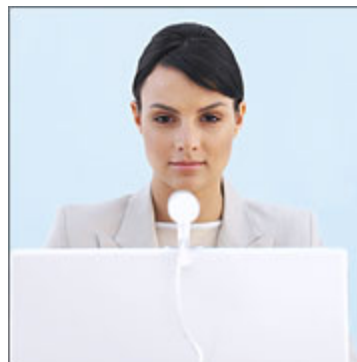
“How the Afghan Girl was Identified by Her Iris Patterns” Read the [story](#)



Login without a password...



Fingerprint scanners on many new laptops, other devices



Face recognition systems now beginning to appear more widely
<http://www.sensiblevision.com/>

Object recognition (in supermarkets)



[LaneHawk by EvolutionRobotics](#)

“A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it...”

Object recognition (*in mobile phones!*)



- This is becoming real:
 - Google Goggles
 - **Lincoln** Microsoft Research
 - Point & Find, Nokia

Special effects: shape capture



The Matrix movies, ESC Entertainment, XYZRGB, NRC

Steve Seitz

Special effects: motion capture



Pirates of the Carribean, Industrial Light and Magic

[Click here for web site](#)

Steve Seitz

Sports



Sportvision first down line
Nice [explanation](http://www.howstuffworks.com) on www.howstuffworks.com

Slide content courtesy of Amnon Shashua

Smart cars

The image is a screenshot of the Mobileye website. At the top, there are two navigation buttons: 'manufacturer products' and 'consumer products'. Below them is the slogan 'Our Vision. Your Safety.' and a top-down view of a car with four yellow cones representing camera fields of view, labeled 'rear looking camera', 'forward looking camera', and 'side looking camera'. Below this are three main product sections: 'EyeQ Vision on a Chip' with an image of a chip, 'Vision Applications' with an image of a pedestrian and text 'Road, Vehicle, Pedestrian Protection and more', and 'AWS Advance Warning System' with an image of a car icon and a distance reading '0.8'. To the right is a 'News' sidebar with two headlines: 'Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System' and 'Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end', followed by an 'all news' link. Below the news is an 'Events' section with two links: 'Mobileye at Equip Auto, Paris, France' and 'Mobileye at SEMA, Las Vegas, NV', followed by a 'read more' link.

manufacturer products consumer products

Our Vision. Your Safety.

rear looking camera forward looking camera side looking camera

EyeQ Vision on a Chip

Vision Applications
Road, Vehicle, Pedestrian Protection and more

AWS Advance Warning System

News

- > Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System
- > Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end
- > all news

Events

- > Mobileye at Equip Auto, Paris, France
- > Mobileye at SEMA, Las Vegas, NV
- > read more

- Mobileye

- Vision systems currently in high-end BMW, GM, Volvo models
- By 2010: 70% of car manufacturers.

Smart cars *are here!*



Nevada embraces the future, approves self-driving cars

By Bill Howard on February 20, 2012 at 9:00 am | [1 Comment](#)



Vision-based interaction (and games)

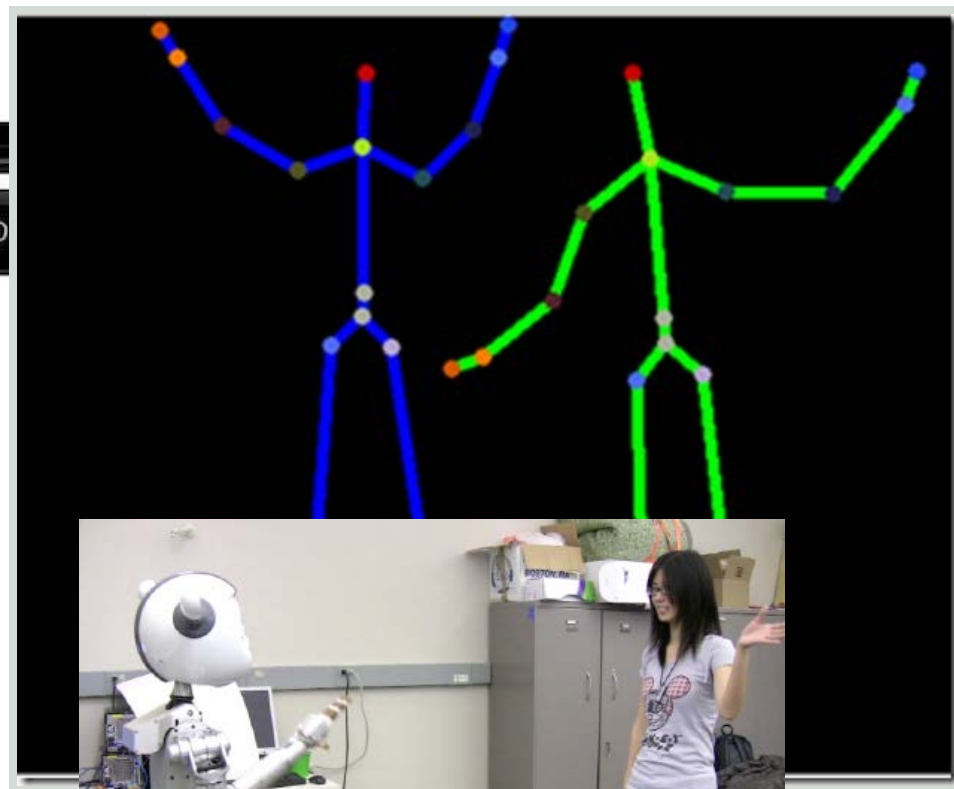


Nintendo Wii has camera-based IR tracking built in.

But the game changer:



KINECT
for  **XBOX 360** Windows



Vision in space



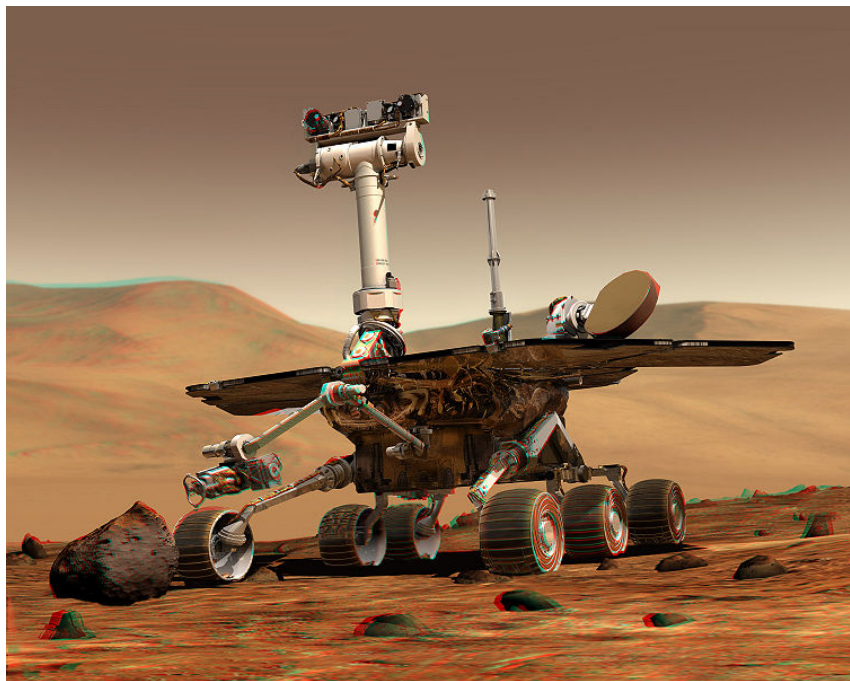
[NASA'S Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “[Computer Vision on Mars](#)” by Matthies et al.

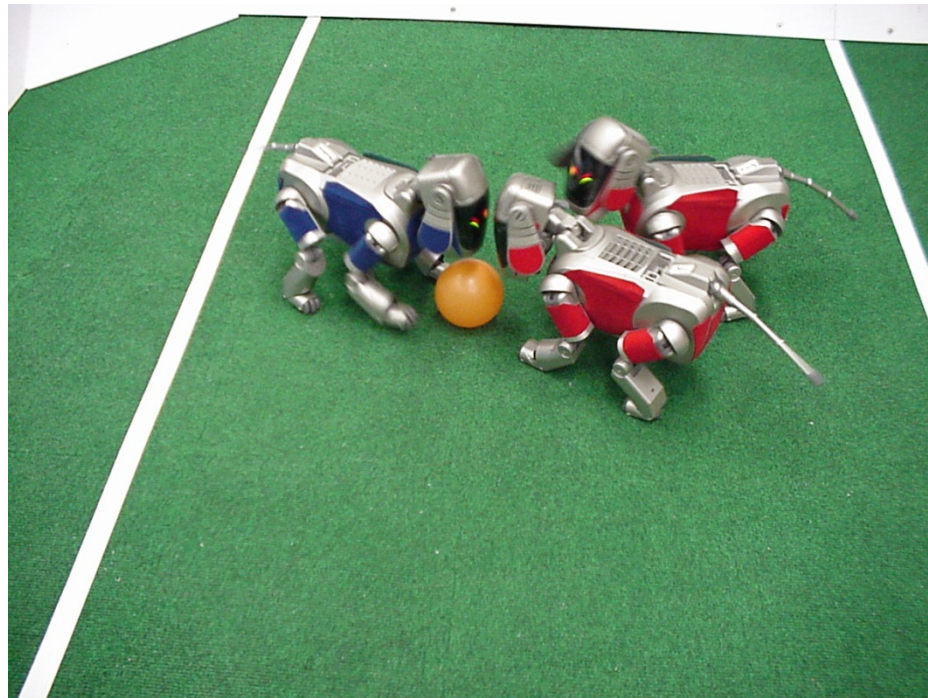
Steve Seitz

Robotics



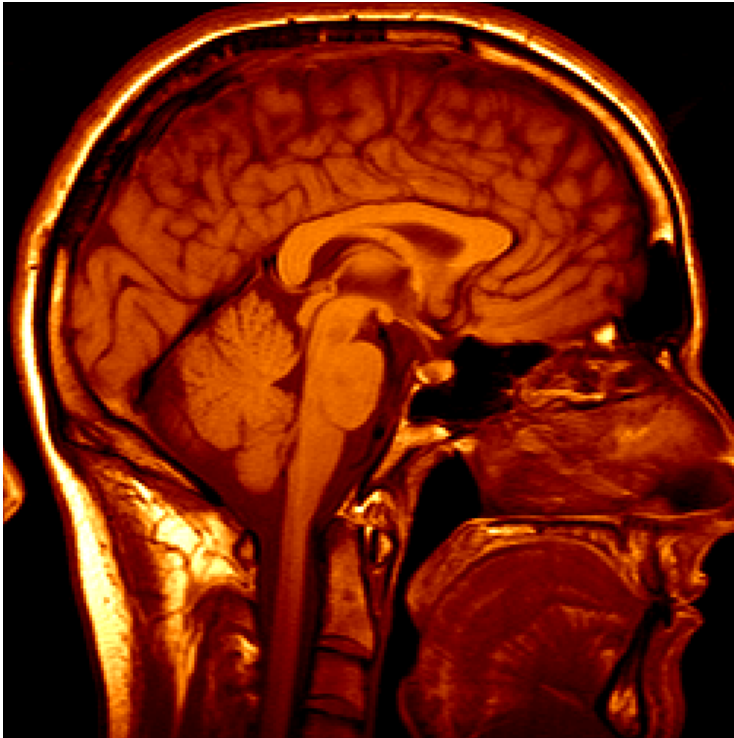
NASA's Mars Spirit Rover

http://en.wikipedia.org/wiki/Spirit_rover



<http://www.robocup.org/>

Medical imaging



3D imaging
MRI, CT



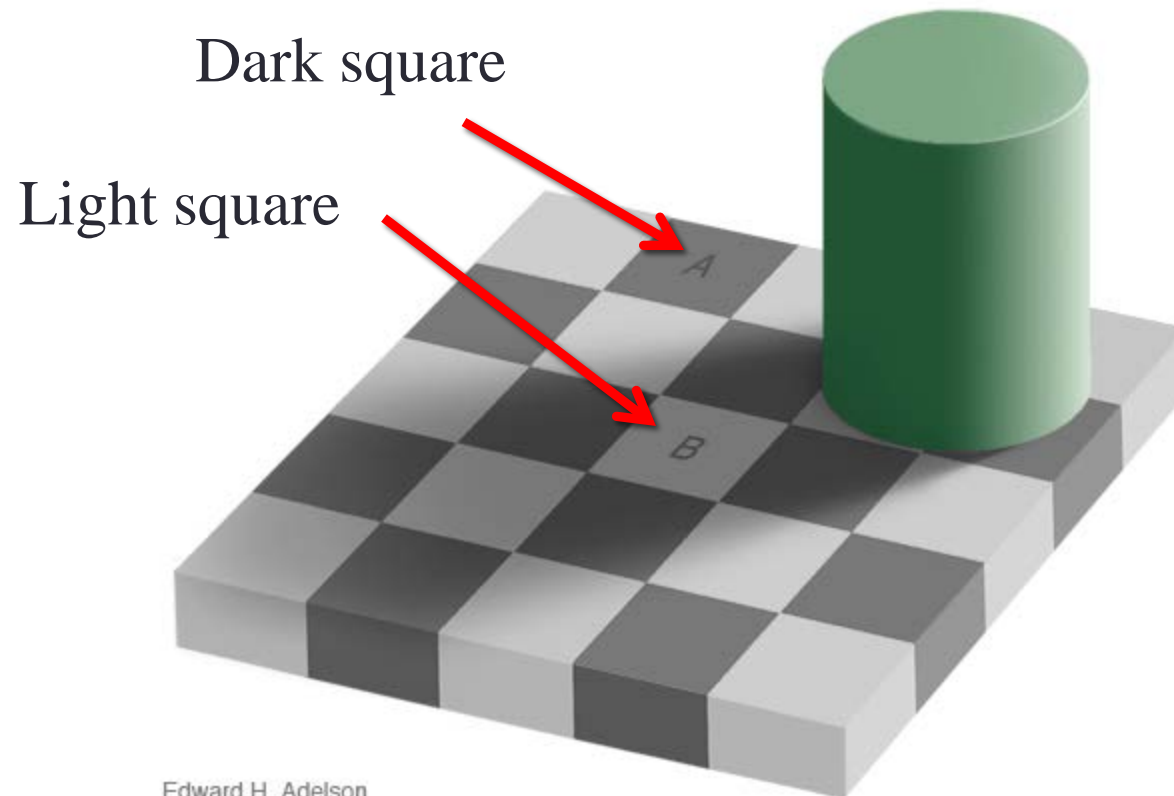
Image guided surgery
[Grimson et al., MIT](#)

Current state of the art

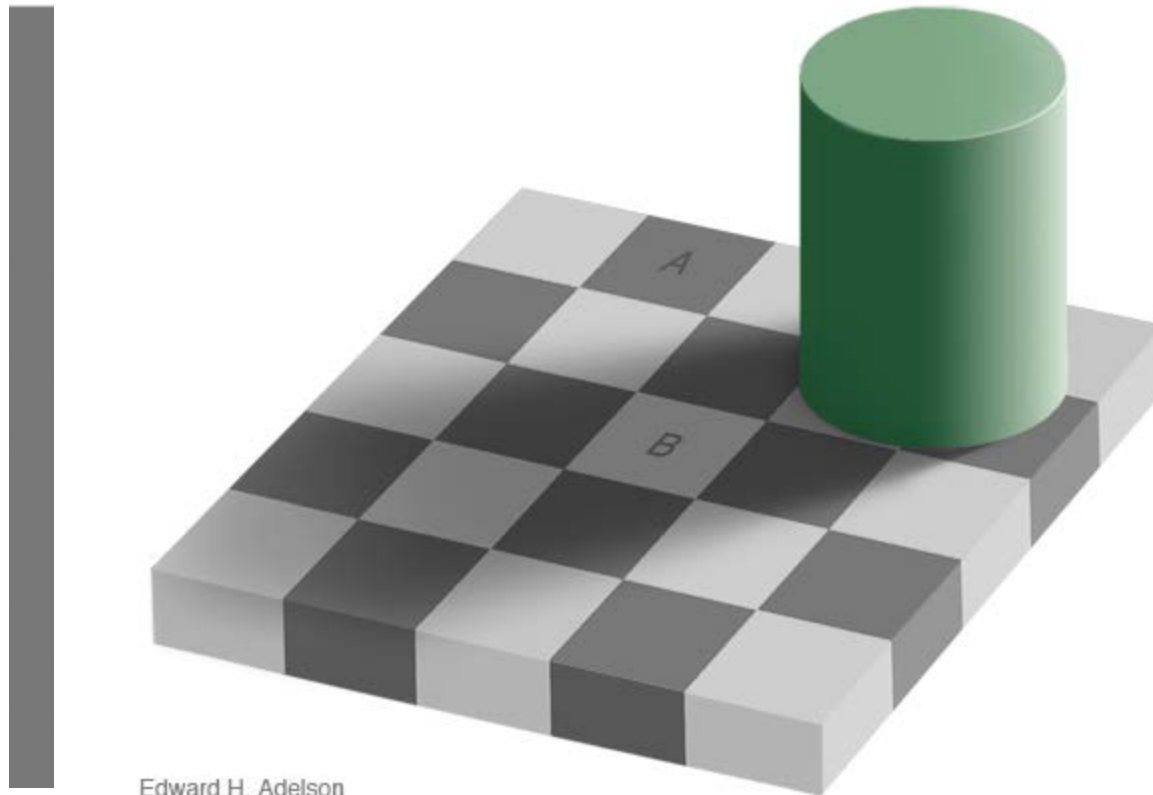
- You just saw examples of current systems.
 - Many of these are less than 5 years old
- This is a very active research area, and rapidly changing
 - Many new apps in the next 5 years
- To learn more about vision applications and companies
 - David Lowe maintains an excellent overview of vision companies
 - <http://www.cs.ubc.ca/spider/lowe/vision.html>

Why is this hard?

Simple scene right?

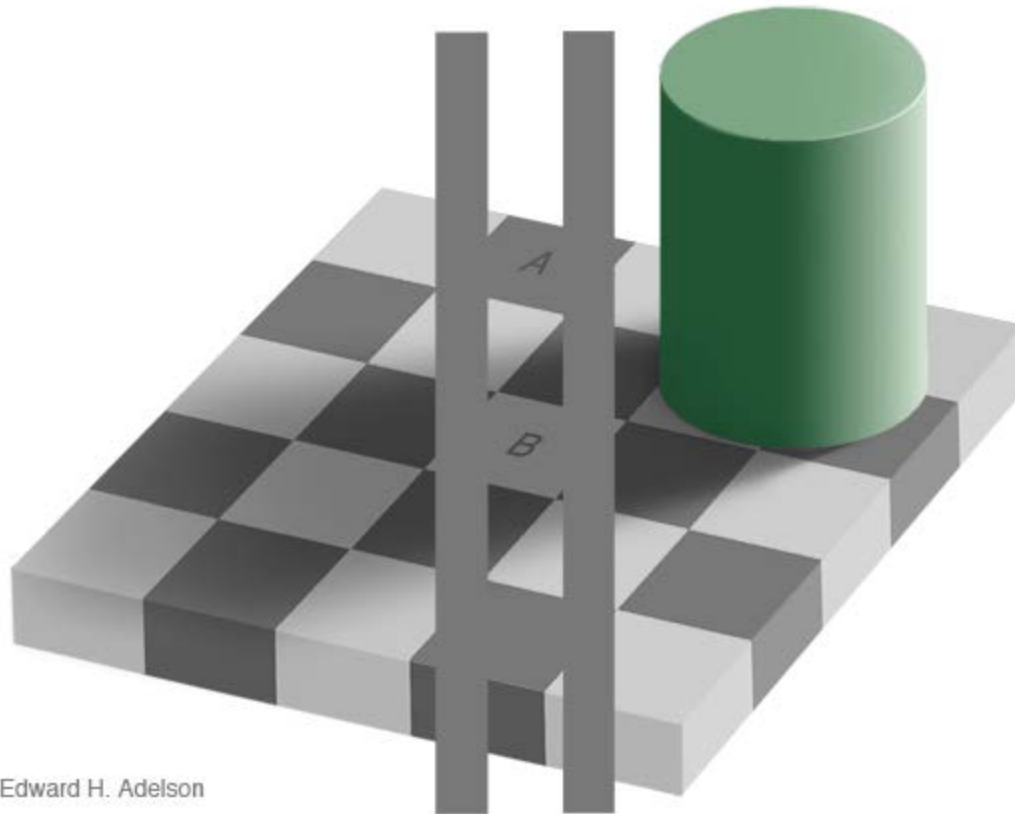


Really?



Edward H. Adelson

Really!

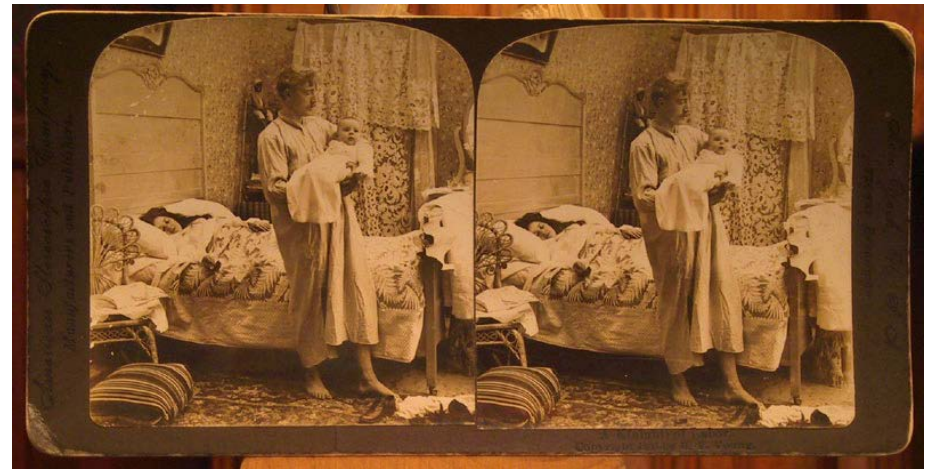


Edward H. Adelson

Vision is NOT Image Processing

- In the previous example, the two squares have exactly the same *measurement* of intensity.
- So, seeing is not the same as measuring properties in the image.
- Rather, “seeing” is building a *percept* of what is in the world based upon the measurements made by an imaging sensor.

Building models from change (1)



Building models from change (1)



Left Image

Michael Black

Building models from change (1)



Right Image

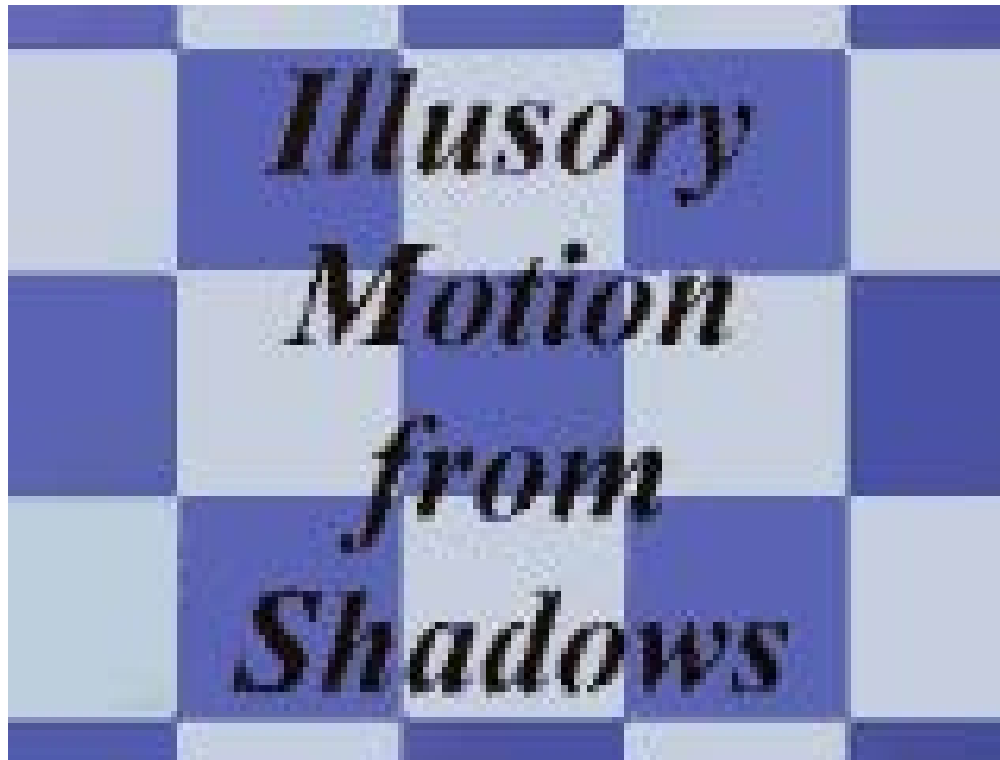
Michael Black

Building models from change (2)



Dan Kersten, <http://vision.psych.umn.edu/users/kersten/kersten-lab/shadows.html>

Building models from change (3)



Dan Kersten, <http://vision.psych.umn.edu/users/kersten/kersten-lab/shadows.html>

Going forward

- For coming lectures:
 - FP: Chapter 4 (all), 5.1, 5.2
- Get yourself Matlab (and/or Python/OpenCV)
- PS0 – yes it's easy but you need to do it and need to submit it in the format requested.